processor (CP) and a slight time delay as illustrated in Fig. 6, some buffering must be included in the system. This is provided by a caching scheme. Such scheme is indicated in greater detail in a copending application, entitled Head and Tail Caching Scheme, Application No. 09/930,804. Referring in detail to Fig. 7, from the communications processor high speed data is coupled through the tail FIFO memory 41 and a multiplexer 42 to the head FIFO memory 43. Data packets will queue up as indicated in Fig. 6 as 1, 2 and 3 and be distributed by the sequential sprinkler engine (SSE) 35 and the read pointer (RP) to the various SEs as discussed. If data comes in at a rate faster than read or outputted to the switching elements fast, and the head FIFO memory 43 fills and the input data will start filling the tail FIFO memory 41. The write pointers and read pointers handle this detail under the control of memory controller 44 which has the WP and RP outputs. It is also coupled to the multiplexer 42. The tail or buffer FIFO 41 will initially keep the head FIFO memory 43 full as it is so-called de-queued (that is as it distributes data packets to the various switching elements). However, if the tail FIFO memory itself becomes full, then the so-called large scale off chip buffer memory 46 is utilized. Here as discussed in the above copending application uniform blocks of data on line 47 are transferred into the memory 46. And the transfer is arranged to be very efficient by use of uniform data block sizes. Finally, when the sudden burst of data packets decreases the traffic manager can de-queue all data from the large scale memory 46 and return to its normal functioning.

The above process is illustrated in Fig. 8 where in step 51 the head FIFO memory is first filled and then in step 52 the tail FIFO memory after the head FIFO overflows. And finally in step 53 the data is stored in the buffer memory until the tail FIFO has space. Then the data is retrieved to the tail buffer and finally written to the head FIFO.

As illustrated in Fig. 4, because of the asynchronous nature of the data inputs to the switching elements and its output as indicated by the time axis reordering may be necessary of the data. In other words, the present invention trades some sacrifices some latency to maintain the highest data rate throughput and enable simple redundancy. Referring to Fig. 9, one reordering technique is illustrated in flow chart form. Here in step 51 each data packet gets a time stamp when it leaves a source communication processor. Then on the output side, when the packets are received by the destination communications processor, they are put into a queue. Each destination CP has a separate queue. As the packets are received, the lowest time

stamp is determined at step 53. A time out period occurs when this system clock reaches the value of the lowest time stamp added to the minimum delay. If this time out period has not yet been released, the system repeats itself as illustrated in step 54. If it has occurred, as shown in step 55, it is now theoretically known that all frames have been received (assuming no other problems) and the packet with the lowest time stamp is placed at the head of the queue. This is just one illustration of reordering and others may be used. However, details of the reordering technique may be found in a copending application titled "Reordering of Sequence Based Packets in a Switching Network:," Application No. 10/044,244.

To provide additional data ports, the switching fabric of the switching elements shown in Fig. 3 is easily scalable or expandable to accommodate greater data input. One technique is a butterfly expansion, illustrated in Fig. 10. Here there are the original SEs, SEO and SEI are so labeled. To expand additional switching elements designated SE2' – SE5' are connected with the designated interconnections that double the amount of input and output ports.

To summarize the operation of the invention, a uniform or variable length data packet is stored in an ingress port at a relatively high data rate and is transmitted to its final destination port on one serial link. Moreover, since the packet is not broken into smaller pieces, where the header becomes a significant part of the data packet, overhead is minimized and the highest data rate is maintained. The switching fabric configuration as shown by the switch elements of Fig. 3 allows for redundancy where, in the case of failure one switch element, another is automatically selected. This is not true of ordinary parallel channel devices as illustrated in Fig. 1. Moreover, additional bandwidth and data input can be provided by adding more switch elements; for example, in a butterfly configuration as illustrated in Fig. 10.

In summary, improved switching apparatus for increasing data rates with limited switching speeds has been provided.

processor (CP) and a slight time delay as illustrated in Fig. 6, some buffering must be included in the system. This is provided by a caching scheme. Such scheme is indicated in greater detail in a copending application, entitled Head and Tail Caching Scheme, ~~Attorney Docket No. 6979/12~~ Application No. 09/930,804. Referring in detail to Fig. 7, from the communications processor high speed data is coupled through the tail FIFO memory 41 and a multiplexer 42 to the head FIFO memory 43. Data packets will queue up as indicated in Fig. 6 as 1, 2 and 3 and be distributed by the sequential sprinkler engine (SSE) 35 and the read pointer (RP) to the various SEs as discussed. If data comes in at a rate faster than read or outputted to the switching elements fast, and the head FIFO memory 43 fills and the input data will start filling the tail FIFO memory 41. The write pointers and read pointers handle this detail under the control of memory controller 44 which has the WP and RP outputs. It is also coupled to the multiplexer 42. The tail or buffer FIFO 41 will initially keep the head FIFO memory 43 full as it is so-called de-queued (that is as it distributes data packets to the various switching elements). However, if the tail FIFO memory itself becomes full, then the so-called large scale off chip buffer memory 46 is utilized. Here as discussed in the above copending application uniform blocks of data on line 47 are transferred into the memory 46. And the transfer is arranged to be very efficient by use of uniform data block sizes. Finally, when the sudden burst of data packets decreases the traffic manager can de-queue all data from the large scale memory 46 and return to its normal functioning.

The above process is illustrated in Fig. 8 where in step 51 the head FIFO memory is first filled and then in step 52 the tail FIFO memory after the head FIFO overflows. And finally in step 53 the data is stored in the buffer memory until the tail FIFO has space. Then the data is retrieved to the tail buffer and finally written to the head FIFO.

As illustrated in Fig. 4, because of the asynchronous nature of the data inputs to the switching elements and its output as indicated by the time axis reordering may be necessary of the data. In other words, the present invention trades some sacrifices some latency to maintain the highest data rate throughput and enable simple redundancy. Referring to Fig. 9, one reordering technique is illustrated in flow chart form. Here in step 51 each data packet gets a time stamp when it leaves a source communication processor. Then on the output side, when the packets are received by the destination communications processor, they are put into a queue. Each destination CP has a separate queue. As the packets are received, the lowest time

stamp is determined at step 53. A time out period occurs when this system clock reaches the value of the lowest time stamp added to the minimum delay. If this time out period has not yet been released, the system repeats itself as illustrated in step 54. If it has occurred, as shown in step 55, it is now theoretically known that all frames have been received (assuming no other problems) and the packet with the lowest time stamp is placed at the head of the queue. This is just one illustration of reordering and others may be used. However, details of the reordering technique may be found in a copending application titled "Reordering of Sequence Based Packets in a Switching Network:," ~~Attorney Docket #06979/08~~ Application No. 10/044,244.

To provide additional data ports, the switching fabric of the switching elements shown in Fig. 3 is easily scalable or expandable to accommodate greater data input. One technique is a butterfly expansion, illustrated in Fig. 10. Here there are the original SEs, SEO and SEI are so labeled. To expand additional switching elements designated SE2' – SE5' are connected with the designated interconnections that double the amount of input and output ports.

To summarize the operation of the invention, a uniform or variable length data packet is stored in an ingress port at a relatively high data rate and is transmitted to its final destination port on one serial link. Moreover, since the packet is not broken into smaller pieces, where the header becomes a significant part of the data packet, overhead is minimized and the highest data rate is maintained. The switching fabric configuration as shown by the switch elements of Fig. 3 allows for redundancy where, in the case of failure one switch element, another is automatically selected. This is not true of ordinary parallel channel devices as illustrated in Fig. 1. Moreover, additional bandwidth and data input can be provided by adding more switch elements; for example, in a butterfly configuration as illustrated in Fig. 10.

In summary, improved switching apparatus for increasing data rates with limited switching speeds has been provided.